UNITED NATIONS

ECONOMIC AND SOCIAL COUNCIL

Distr.
LIMITED

E/ECA/PSD.3/43
3 October 1983

Original: ENGLISH

ECONOMIC COMMISSION FOR AFRICA

Third Session of the Joint Conference
of African Planners, Statisticians
Demographers

Addis Ababa, 5-14 March 1984

STATISTICAL COMPUTING PROJECT / UNDP-SPONSORED REGIONAL PROJECT
IN THE ECA REGION

TABLE OF CONTENTS

		Paragraphs
I.	INTRODUCTION	1 - 3
II.	OBJECTIVES OF THE PROJECT	4 - 7
III.	INSTITUTIONAL FRAMEWORK	8 - 11
IV.	STRATEGY OF ACTIVITIES	12 - 15
V.	DATA EDITING	16 - 19
VI.	TABLE GENERATION	20 - 23
VII.	STATISTICAL DATA BASE MANAGEMENT (SDBMS)	24 - 27
VIII.	META-INFORMATION SYSTEM	28 - 31
IX.	LONG-TERM OBJECTIVES OF THE PROJECT	32 - 35
X.	TRANSFER OF OUTPUT OF THE SCP TO DEVELOPING COUNTRIES	36 - 40

OUTPUTS OF THE STATISTICAL COMPUTING PROJECT ENVISAGED
TO BE COMPLETED BY THE END OF 1984

I. INTRODUCTION

1. The intercountry project on the use of computers for statistical purposes and the design and development of automated statistical information systems was established in 1968 by participating Governments and the United Nations Development Programme (UNDP).

2. In recent years, the central statistical offices of the Economic Commission for Europe (ECE) member countries have engaged in an exchange of experience in the use of computers and in the design and development of statistical information systems. The exchange has taken place mainly at a series of seminars held annually on Integrated Statistical Information Systems (ISIS) at the Computing Research Centre, Bratislava, Czechoslovakia. The practical contribution to the seminars has shown that work is proceeding in a number of ECE countries on the same or similar problems. With the assistance of the ECE and of some Eastern countries, UNDP agreed on a regional project on the exchange of experience and information between interested European countries.

3. The following countries are participating in the project: Bulgaria, Canada, Czechoslovakia, Denmark, Finland, France, German Democratic Republic, Greece, Hungary, Ireland, Italy, Netherlands, Poland, Romania, Spain, Sweden, Turkey, United Kingdom and Yugoslavia.

II. OBJECTIVES OF THE PROJECT

4. The programme is aimed to increase the efficiency of national statistical offices with a better use of computers, in order that policy-makers and planners have more accurate and more up-to-date data. With a division of labour in programming and in the development of software products between national statistical offices, the skill and experience of highly developed countries will be transferred to countries with lesser developed computerization of statistical services.

5. The project has two primary objectives: (a) joint development of urgently needed software for higher efficiency of national statistical services and (b) particular benefit for countries eligible for UNDP assistance in gaining experience and training while co-operating with developed countries not eligible for UNDP assistance in the development of software products.

6. The immediate objectives of the project are to establish a programme on the use of computers for statistical purposes and for design and development of automated statistical information systems. These will include the organization of joint work between national statistical offices of participating countries on the design of software products for every-day work, including applied research, development and testing of software products and related tools specially designed for use in practical statistical work. It is aimed to provide countries with limited experience in software development, with an opportunity to become acquainted with advanced methods in this field of the statistical data processing and to receive the necessary training.

7. It is foreseen that the software products, developed by participating countries within the framework of the programme, will be made largely available to developing countries. Thus, the results of the project, as well as the methodology developed, will be useful not only to the participating countries but also to those outside the ECE region - in particular, to the developing countries.

III. INSTITUTIONAL FRAMEWORK

8. The Project is being carried out jointly by the statistical offices of the participating Governments, UNECE and the ECE secretariat. It is administered by a Steering Committee assisted by the Project Co-ordinator.

9. The Steering Committee consists of representatives of the participating countries. It forward to take decisions regarding the programme. The programme is conducted by the Steering Committee, responsible for planning and implementing the Project, taking into account the procedures and requirements of the UNECE and the views and comments of the Conference of European Statisticians. The Steering Committee operates according to its Rules of Procedure. Between the sessions of the Steering Committee, its Bureau gives support to the Project Co-ordinator, who is responsible for the implementation of the Project.

10. The Steering Committee, at the time of its first session in Geneva in December 1970, agreed to create four Joint Groups to work on selected topics and research tasks. These Joint Groups are:

(a) Table Generation, with the participation of Czechoslovakia, Denmark (lead country), Finland, German Democratic Republic, Poland, Romania, Turkey and the United Kingdom;

(b) Data Editing, with the participation of Bulgaria, France, Greece, Hungary (lead country), Ireland, Spain, Sweden, the United Kingdom and Yugoslavia;

(c) Metainformation System, with the participation of: Bulgaria, Canada, Czechoslovakia (lead country), Netherlands and Israel;

(d) A relational data base for Statistical Data Base Management (SDBM), with the participation of Bulgaria, Canada, German Democratic Republic, Hungary, Italy, Sweden (lead country) and Yugoslavia.

11. The practical work has been undertaken by teams of participating statistical offices working together in joint groups, but some tasks have been performed by a single office in accordance with the agreed programme of work for the respective joint groups.

IV. SUMMARY OF ACTIVITIES

12. In conformity with the decisions of the Steering Committee and the objectives of the Project, joint work has been undertaken and will be continued in the design of software products of practical use, including applied research, and on the development and testing of software products. The Joint Groups have organized the practical work between the national statistical offices of participating countries by bringing together resources from these statistical offices to work on problems of general interest.

13. Joint work has been performed within the framework of the Joint Groups, in accordance with their work plans for the period between two sessions of the Steering Committee. The activities were reviewed at meetings of the Joint Groups and by the Steering Committee. A regular review of the joint work has been, and will be made at Joint Group meetings together with updating of their programme of work.

14. In 1981 the Joint Groups started their joint work. The general aim of the Joint Groups during this phase was to produce working papers for system specifications and model materials for the organization of recapitulations at a later date. The working papers distributed to the participating countries were discussed during the meeting of the Joint Groups in 1981.

15. In 1982, considerable progress has been achieved with the joint work carried out as follows. The working papers together with the work plan for the project were reviewed at the Joint Groups' meeting held by the Steering Committee in November 1982. In addition to the Joint Groups' meetings, five _____ sessions of experts were held to deal with questions delegated by the Joint Groups to small task forces.

V. DATA EDITING

16. As statisticians we know data editing is essential to ensure consistency and quality of the output produced in statistical offices. It has a heavy impact on the workload of many time-consuming statistical operations made in order to improve the timeliness of putting socio-economic information at the disposal of planners and policy makers. Whether there is little outside support from the theoretical viewpoint, and hardly any commercial software available, data editing for statistical purposes requires specific solutions found by experts from national statistical offices.

17. As regards the evaluation of data editing methods actually applied in national statistical offices of the ECE region, an analysis has concentrated on the hardware and software aspects of interactive editing systems. Methodological document was prepared serving as a basis for further work on the data quality aspect in connection with generalized editing systems.

18. The generalized data editing system called DIA was developed in Hungary. Following several visits of statistical officers to their installations, experiences, a new version of DIA was prepared including the necessary system documentation which is now available. The DIA system has been transferred to other statistical offices. In particular the work to make the DIA software system interactively usable was discussed by visitors of these offices. In order to make the DIA system portable as much as possible, a version DIA/Version 2 with PL/1(F) compiler together with its machine-readable documentation has been started.

19. A new field of activities was assigned by the Joint Groups to the survey processing. After a review of existing approaches, functional and development aspects were identified to serve as a basis for further work in the field of the generalized survey processing system (GSPS). It is envisaged that functional specifications of GSPS should be accomplished within the limits of the ECE. The specifications should focus on the selection criteria and interface definitions when existing programs and software elements are analysed to form a subsystem of GSPS. It is believed that this approach would help statistical services to clarify questions in designing survey processing.

E/ECA/PSD.3/43
Page 4

VI. [illegible]

20. In 1971, starting with a detailed list of functional characteristics, the Joint Group developed the main functional and structural concept of a new table processing system called INTERTAB. At the same time, existing table processing packages as known to the Group were scrutinized with a view to possible modifications to fulfil the requirements that had been worked out. Even though numerous tabulation programmes were available, no existing system was found an appropriate candidate for such modifications. At a later date, the TAU system, being developed by the Office for National Census and Surveys, United Kingdom, had been also evaluated but found to be unsuitable for modifications in the sense of specific functional characteristics of INTERTAB.

21. Based on the assessment of the existing facilities and a higher level of table production was desired by developing a new system, such as INTERTAB, the Group devoted its activities to the development of INTERTAB. Efforts were focused on the system architecture, the counting techniques and the user language of INTERTAB. Within papers on these system parts have been produced and discussed.

22. Having finalized the general system architecture together with the functional characteristics of INTERTAB, the Group turned to the implementation. In order to speed up the implementation of INTERTAB, a series of expert group meetings took place during the implementation. It was agreed that the implementation of INTERTAB will take place in four development centres and one test centre. Each centre is fully responsible for performing the task assigned in accordance with the time-table, the procedures and the standard's agreed upon. The actual coding work will be in in June 1982 and the first system test is envisaged to be performed in May 1984. INTERTAB (Version 1) which was specified to run on IBM or IBM-compatible computers, is planned to be completed by November 1984.

23. In parallel to INTERTAB implementation, another table generator system, called TAU, is presently under development in the United Kingdom for installation on ICL computers. The present version of TAU was transferred to Poland and Romania on testing terms on their ICL computers.

VII. STATISTICAL DATA BASE MANAGEMENT (SDBM)

24. It was recommended there that improving the access to the statistical data stocks, which are already stored on magnetic media, should have primary importance. For national statistical offices, the issue of appropriate data base management system is of importance in view of their large and ever-increasing data stock. The relational approach appears adequate since data of a statistical surveys have well-defined relations to each other.

25. The Joint Group on SDBM decided on further development of the relational data base management system, called RAPID and developed by Statistics Canada, with a special view to creating interfaces to existing software systems that serve other main functions of statistical work. The joint software development concept was launched and planning of new SDBM software modules was the tasks of experts who had [illegible] practice in planning and testing of new software pieces. The successful implementation of the first new module proved the effectiveness and usefulness of the small programming team effort, where small portions of coding were allocated to individual programmers, while the planning and the final testing of the module were performed jointly.

26. Transfer of statistical software packages and its effective implementation [illegible] in Hungary and in the German Democratic Republic. The installation in Hungary, which at existing [illegible] Statistics, was performed in three stages of [illegible] with a technical environment changed significantly between [illegible] operating systems. The installation in [illegible] Republic is a pilot application in international statistics and particularly one meant to explore the conditions arising under different [illegible] similar to IBM/360 or IBM/370.

27. Review of numerous [illegible] in Europe and North America for their encountered statistical [illegible] and their [illegible] statistical data base management tasks. [illegible] installations, Bulgaria, Italy and Yugoslavia [illegible] implement UNIDO. The planned activities [illegible] the creation of [illegible] installations will cover both surveys of the existing UNIDO installations [illegible] machine portable basic machines (host computer suite) (i.e. IBM mainframes). In addition, a comparative evaluation of selected [illegible] management systems in the light of [illegible] statistical applications will be included in the work programme for 1984.

VIII. [illegible]

28. The basic design of the [illegible] system specifications for the area of the statistical information system. The increasing [illegible] into stocks and the ever [illegible] economic information call for an efficient system of information for the statistical system as a whole. Design of a general data [illegible] information system together with its logical specifications is [illegible] research field.

29. The actual state of meta-information systems (METIS) development was investigated by a survey conducted within the participating countries in 1981 and the analysis and evaluation report was prepared. Starting [illegible] elaborated for some of the participating countries' state statistics, a METIS conceptual report was produced. During this period the Group's activity focused on rules actually applied for statistical [illegible] and experiences gained from their use.

30. The Group set up the specifications of objects dealt within METIS, methods for formalizing METIS object descriptions and ways of [illegible] descriptions on software and hardware requirements for meta-information systems. As regards the METIS objects functional specifications were separately defined for the end-user, the management and the information system designer's viewpoint. Work on the end-user's and management's functional aspects have been completed. A synthesis of the specifications [illegible] functional specifications has not been completed in order to yield an integrated picture of the requirements.

31. Formalized description of the METIS subsystems elements to [illegible] for the objects of the statistical information system. In this respect, specification of the METIS software functions will [illegible] particular for implementation and maintenance of a meta-information system where existing software shall be scrutinized for eventual employment. A comprehensive final report on the METIS field is planned to be completed by end of 1984.

IV. LONG-TERM OBJECTIVES OF THE PROJECT

32. [illegible] actual state of the art in the field of data edition, survey processing, [illegible] information systems and data base management in the statistical offices participating in the ECA constitutes an essential prerequisite for attaining the ultimate objective of the software development. In this respect, the Project envisages to elaborate reports which will be [illegible] manuals for future guidance.

33. [illegible] statistical software development. Many aspects in both areas recently [illegible] have closely influenced the actual software basis in the field of data edition, tabulation and statistical data base management. In this connexion, discussions of terms have to be conducted, constituting a basis for the improvement of such a [illegible] decision which is of great importance for future joint conceptual work on software development.

34. It has to be particularly emphasized that the decision of not limiting the work to the development of new software was especially valuable and realistic. In harmony with this Project objective, the analysis of existing software packages will provide a sound basis for maximum exploitation of existing software systems. This approach has now to be developed in several interfaces and new modules for linking the SCO [illegible] must be numerous existing statistical software packages. Creating new software systems suitable to other than the basic machines involves the essential reason of its application. This concept has already been implemented in many areas of the Project.

35. The Steering Committee at the [illegible] Statistical Data Processing Methodology should be published. The [illegible] is time to represent the first volume of a [illegible] series of such [illegible] currents. It will cover the technical [illegible] of statistical data processing, but it will not have a total coverage as the [illegible] of publications on statistical computing in general.

V. TRANSFER OF OUTPUT OF THE SCO TO DEVELOPING COUNTRIES

36. The Project has always regarded the first [illegible] as the executing agency of the Project should not forever [illegible] interaction in order to promote the actual transfer [illegible] of the SCO to developing countries, in harmony with the decision of the Steering Committee in this respect. It should be emphasized that the Regional Commissions and in particular the statistical divisions/secretariats are the main instrument for international co-operation during the course of transfer of the SCO output. However, the developing countries should receive the required assistance of the United Nations system in the effort of obtaining the SCO output, including substantial assistance from the UNDP.

37. In respect to the actual work towards this direction, the following should be mentioned. A scientific meeting on "Statistical Software for Developing Countries" will be organized, which will aim at bringing some output of the SCP to the attention of developing countries present at the ISI Conference (September 1983, in Madrid). Consultation should be initiated with regional commissions on selecting potential candidate statistical offices for transferring know-how of the SCP. A Workshop on the SCP is envisaged to be organized in Europe, in mid-1984. It is aimed at bringing more detailed information and experience or usage of software packages developed within the framework of the Project to experts of statistical services of developing countries. At that time FDP experts of the ECE region can provide overall information on the experience of the international co-operation of system development in connection with the envisaged output of the Project (See Annex I).

38. To properly software products and related tools of the SCP to users in developing countries, it is necessary to do much more than develop and distribute systems. From a realistic point of view user training will also have to be carried out for and, depending on the properties of an individual package, personal installation support will have to be casually provided. An appropriate organization must be established and maintained which can provide adequate support services to software users.

39. While most developing countries show a considerable degree of heterogeneity with respect to their statistical data processing, it should be noted that certain common differences can be identified as compared with the majority of the participating ECE countries of the SCP. Therefore, it is fundamental to pay attention to those peculiarities of developing countries with respect to the general scope of efficient support in the field of statistical computing.

40. In the light of the indicated interest in the SCP, the Economic Commission for Europe has the view that it would be useful to specify a co-operative plan of action specifically oriented towards the transfer of the know-how of the Project to developing countries. In this respect the regional commissions are invited to co-operate with ECE to make use of the SCP output in statistical organizations in their region.

No.	Activities	Output	Scheduled completion
1	Development of new software elements for ICLIP: User operators	Implementation & test of 'user operators' – software design, coding and installation – documentation	31 1984 4 1984
2	Consolidation of ICLIP application for Census data research, Hungary	Final report on application Census Data Base with ICLIP	15 1984
3	Implementation of ICLIP on selected computer systems	Implementation of ICLIP on ICL 2900 system Implementation of ICLIP on ICL 1900 system	30 1984 15 1984
4	Comparative analysis of software tools used for statistical applications	Report on the comparative analysis	30 1984
5	Consultancy and advice to ICLIP users	ICLIP Guidelines for users User's manuals, on-line documentation	30 1983 30 1984
6	Implementation of ICLIP version 1	User's software package for ICLIP version 1 Full set of user documentation on ICLIP version 1 Internal documentation of ICLIP V1	1 1984 1 1984 30 1984
7	Implementation of ICLIP distribution	Implementation of ICLIP on ICL mainframes Implementation of ICLIP outside the United Kingdom System documentation on ICLIP	30 1983 20 1984 15 1984
8	Comparative analysis of selected table generating systems	Report on the comparative feature analysis	20 1984

N.			
9			
10			
1			
2			
3			
4			
5			
6			